# iGEM 2008: Engineering *E. coli* to Overexpress Folate

## Background Information

Despite decades of research and characterization of biological pathways, interactions, and genetic information, engineering even the simplest biological system remains expensive, complicated, and, worst of all, often unpredictable. It can be hard to tell whether an experiment failed because there was an unknown interaction between enzymes, or because the conditions were not exactly right for the strain, or perhaps it was just because you forgot to knock on wood and spin three times before incubating your samples. Synthetic biology aims to reduce the level of uncertainty in engineering new biological systems by creating standard parts, known as "BioBrick™ parts," which exist as DNA sequences known to perform specific functions. The hope is that eventually, each BioBrick™ part will be so well characterized that bioengineers and chemical engineers of the future will no longer need to worry about the compatibility of basic components and can instead focus their energies on building more complex systems (Endy 2005).

The purpose of the International Genetically Engineered Machine (iGEM) competition is to apply the principles of synthetic biology and design a synthetic biological system from standardized parts, and in the process support or disprove the claim that biology can be standardized. In the process, undergraduates get the opportunity to test the effectiveness of existing standard parts, to design a potentially useful biological system, and to add new parts from their own project to the pool. The competition culminates at an annual Jamboree held at MIT in November, for which each team will make a poster and also give a half-hour presentation on their project.

For our iGEM team, the project has three principal components: selecting and designing the biological system, doing the wet-lab construction and characterization of the system, and presenting the results at the 2008 iGEM Jamboree. Since the wet-lab construction and cloning is the most time-intensive step, we wanted to have the entire summer free to work on step two. As such, the team began meeting via conference call in April, during which we discussed our overall goals, what system we wanted to design, and then the breakdown of that larger system into individual projects. From these meetings, in addition to online research, we were able to decide on a project and assign individual components prior to arriving on Caltech's campus.

This year, Caltech's iGEM team has decided to create the ultimate gut flora out of an existing probiotic strain of *Escherichia coli*. After all, $10^{14}$ bacteria already live in the large intestine, breaking down energy sources and producing vitamins that humans would otherwise be unable to access (Wikipedia n.d.). Our engineered strain would have four possible states, each of which would allow it to perform a different useful function for its human host. In its undifferentiated state, the bacteria would by default produce a constant stream of folate, an essential vitamin for human survival. Using a stochastic epigenetic switch, each individual bacterium would be able to differentiate into three other states: one state would quickly break down lactose as soon as it entered the gut, another state would produce hydrogen peroxide upon detecting pathogens, and the third state would randomly release pathogen-specific phages. The rapid lactose breakdown state is intended to decrease the effects of lactose intolerance, which is usually caused by other strains of gut bacteria fermenting the sugar and releasing methane gas. The latter two states would maintain a constant defense against pathogens since one is randomly releasing a

pathogen-specific phage, while the other uses specific pathogen detection before releasing a very nonspecific killing reagent. Since both lactose breakdown and releasing hydrogen peroxide involve lysing/killing the individual cell, the folate made prior to differentiation would be released at that time. The hope is that individual cells would lyse often enough such that there would also be a constant supply of bioavailable folate for the gut to absorb.

My individual project is to engineer the folate overproduction in *E. coli*. Folate, the generic term for the various forms of Vitamin B9, is an essential vitamin because it is heavily involved in amino acid synthesis as well as single-carbon transfer reactions. Folate deficiencies in women can result in birth defects such as neural tube defects and other spinal cord malformations. As important as folate is, humans are unable to produce folate, and so must obtain it from eating foods such as green leafy vegetables or folate-fortified cereals (W. Sybesma 2003). An engineered strain of bacteria that would constantly release folate into the gut would reduce the need to fortify breads and cereals with folate, as well as reduce folate-related birth defects in regions with little access to folate-containing foods. In addition to all the reasons stated above, folate is an ideal vitamin to be produced in the gut because it has been shown to be absorbed in physiologically relevant quantities in the large intestine (Asrar and O'Connor 2005). This is not the case for many vitamins (Asrar and O'Connor 2005).

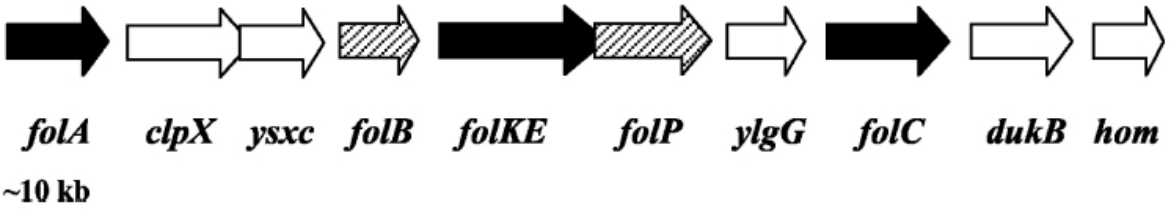


**FIGURE 1: THE *L.LACTIS* FOLATE GENE CLUSTER (SYBESMA, ET AL. JUNE 2003). EACH ARROW REPRESENTS THE RELATIVE SIZE OF EACH GENE. BLACK ARROWS ARE GENES WHICH HAVE BEEN TESTED FOR THEIR EFFECT ON FOLATE SYNTHESIS, SHADED ARROWS ARE GENES WHICH ARE ALSO INVOLVED IN FOLATE BIOSYNTHESIS, AND WHITE ARROWS ARE GENES WHICH ARE NOT INVOLVED IN FOLATE BIOSYNTHESIS (SYBESMA, ET AL. JUNE 2003).**

Although folate is naturally produced in *E.coli,* the folate biosynthesis pathway in the bacteria *Lactococcus lactis* has been more heavily characterized and studied. There are six major enzymes involved in folate synthesis, which, in *L.lactis*, are contained in five genes: *folB, folKE, folP, folC,* and *folA* (Sybesma, et al. June 2003). The first four, which we have chosen to focus on, are located in a gene cluster approximately 4.4kb long. We've chosen not to focus on *folA* for the time being because *folA* encodes an enzyme which turns ones form of folate (dihydrofolate) into another form of folate(tetrahydrofolate). Since our assay would detect both types of folate as part of the total folate production, *folA* was not a prime target for overexpression of folate. In previous studies, this folate gene cluster has been successfully transformed into the folate-consuming bacteria *L.gasseri*, turning the bacteria in to folate-producers (Wegkamp, Starrenburgl, et al. May 2004). Therefore, we have chosen to also use the folate operon from *L.lactis*, which also offers the additional benefit of removing the operon from its natural regulatory context.

Our strategy is to clone the entire folate operon from the *L.lactis* genome and to transform the entire operon into *E.coli*. However, because we are unsure of whether or not the ribosomal binding sites (RBS) within the *L.lactis* operon would work in *E.coli*, we are also going to unpack the operon by cloning each of the four genes individually, placing them behind *E.coli* RBSs, and then recombining them into a single empty BioBricks™ plasmid. In addition to the main folate operon, we will also be experimenting with overexpression of the para-aminobenzoic

acid (pABA) synthesis pathway from chorismate. Wegkamp *et al* have shown that in order to increase overall total levels of folate, both the pABA synthesis genes and certain folate production genes need to be overexpressed (Wegkamp et al., Apr. 2007). The pABA pathway involves three genes, *pabA, pabB,* and *pabC* – though in *L.lactis, pabB* is actually a fusion gene encoding for both *pabB* and *pabC* (ibid) .

## Overall System Design

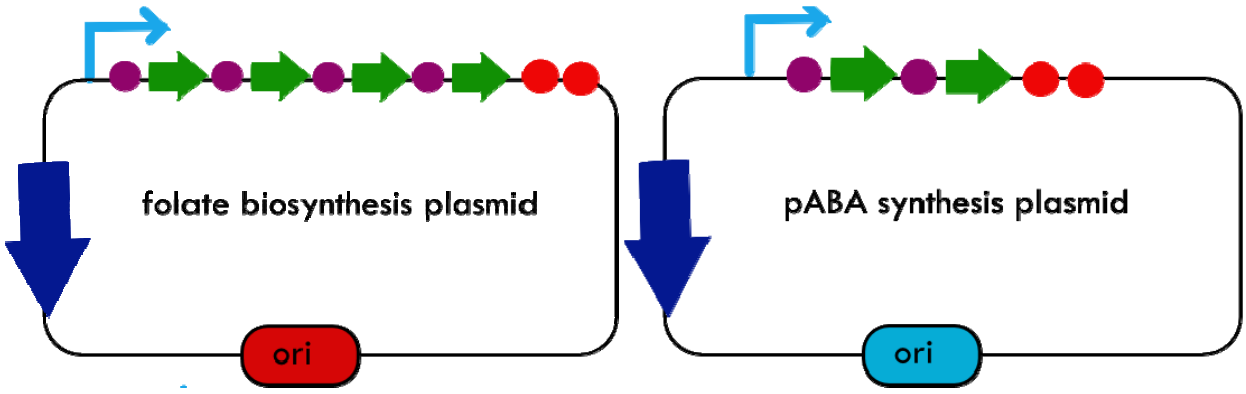


**FIGURE 2: OVERALL SYSTEM DESIGN FOR OVEREXPRESSION OF FOLATE GENES IN *E.COLI***

The overall system design for testing folate production in *E.coli* is to construct two plasmids (Fig.2) – one for the folate biosynthesis pathway, and one for the pABA synthesis pathway. In addition to ensuring that the plasmids are complementary, each plasmid would need to contain a different variable copy origin of replication, which would be low copy by default, but can be switched to high copy via the addition of Isopropyl β-D-1-thiogalactopyranoside **(**IPTG)to the media. This will allow us to test overexpression of each plasmid separately. In addition, each plasmid will contain a constitutive promoter, since we would want folate to be produced constantly. The purple dots represent ribosomal binding sites (RBS), followed by the gene (green arrow), and eventually terminating in a double stop (TAATAA) sequence, as regulated by the Registry of Standard Biological Parts.

## Current Progress

During the first two weeks of my research, I have been cloning all seven genes (the entire operon, *folB, folKE, folP, folC, pabA,* and *pabB)* from the *L.lactis* genome using polymerase chain reaction (PCR) and with varying success. This initial cloning step involves designing and ordering primers for PCR, running the PCRs, and then running the resulting DNA on gels to determine the success of the PCR. For the genes that I have had success in cloning and isolating (*folB, folKE, pabA,* and *pabB*), I have gone on to digest with the restriction enzymes XbaI and SpeI, which is necessary for insertion into the BioBricks plasmids. After the inserts and the vectors were cut, they were ligated together and transformed into competent *E. coli*. These transformed bacteria were incubated and then plated onto LB-Ampicillin agarose plates, which would selectively kill any bacteria which did not receive the plasmid. The success of the transformation can then be evaluated by the number of colonies growing on the plates after overnight incubation.

## Goals for the Coming Month

Six of the eight transformations were successful, though the two that failed (*pabA*+B0034, *pabB*+ B0034) were with B0034 which is the RBS that needs to be added in front of the genes. For the successful transformations, we are growing up those individual colonies overnight, and then we can begin attempts to ligate several parts together into a single plasmid, and to select and add a promoter. As shown in Figure 3, we are currently only beginning Step 1 of plasmid construction. In addition, more cloning needs to be done to isolate the genes which were not successfully cloned the first time (entire folate operon, *folP,folC*). Hopefully we can successfully clone *folC* soon so that site-directed mutagenesis can be done to modify the two *EcoRI* sites within *folC*.

Eventually, once most of the parts have been successfully cloned and inserted into the correct plasmids, we will be detecting folate production, and thus the relative success of our engineering efforts, via a microbiological assay involving the folate-dependent strain *Enterococcus hirae* (Horne and Patterson 1988). This assay involves first the characterization of a standard growth curve of *E. hirae* given known quantities of folate present in the growth media. Once the standard curve has been established, then experimental growth levels, as quantified by spectrophotometry, can be interpolated. PABA concentrations will be measured via high performance liquid chromatography (HPLC) (Wegkamp, Oorschot, et al. Apr. 2007).

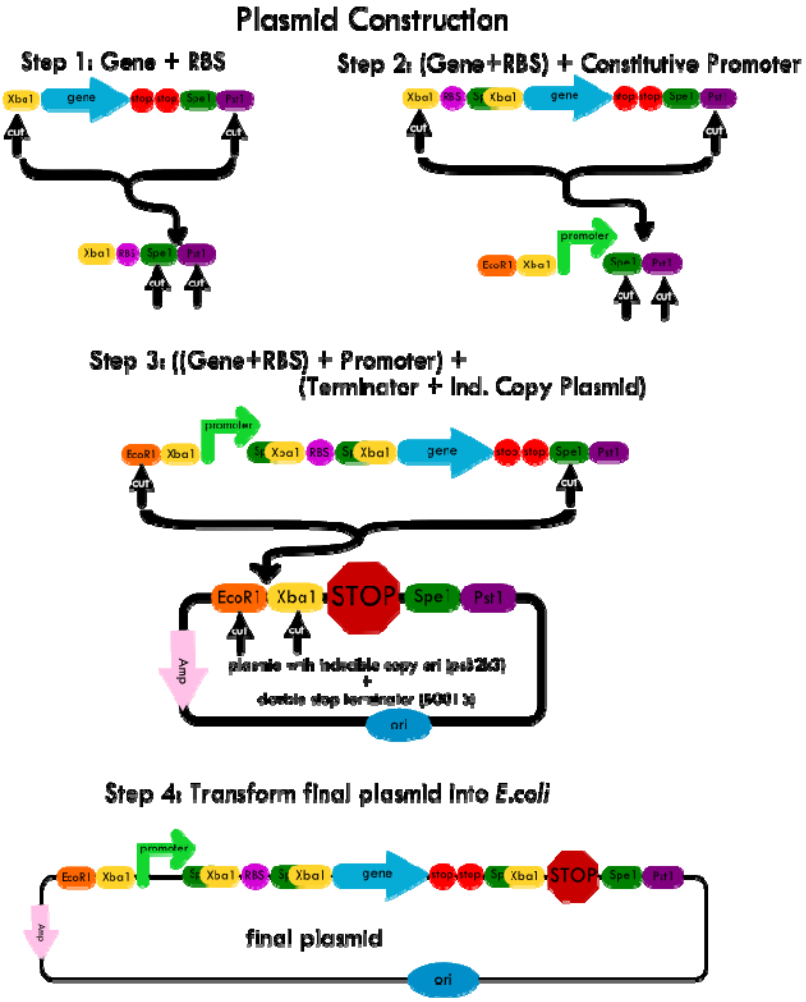


**FIGURE 3: PLASMID CONSTRUCTION STEPS. THERE ARE FOUR RESTRICTION ENZYMES INVOLVED, EACH WITH THEIR OWN UNIQUE RESTICTION SITES. EACH RESTRICTION DIGEST IS COMPATIBLE TO ITSELF, AND THE STICKY ENDS OF XBA1 AND SPE1 CAN BE LIGATED TOGETHER INTO AN IRREVERSIBLE MIXED SITE.**

## CHALLENGES & PROBLEMS

Currently we are having some problems cloning two of the genes, *folC* and *folP*, in addition to also being unable to clone the entire folate operon. To solve this we are retrying the PCR and modifying the melting temperature, annealing temperature, and extension times. We will run a gradient PCR of different annealing temperatures for both *folP* and *folC* to try and screen for the best temperature. Then we would repeat the PCR with just that one temperature. Once that works, we will also have challenges with the site-directed mutagenesis of *folC*. Another challenge coming up is trying out the folate assay and working through the procedure for that.

## RESOURCES REQUIRED

Currently the resources that I am using are standard molecular biology lab materials such as agar plates, the PCR machine, Eppendorf tubes, etc. Specially ordered resources include primers designed for cloning specific genes, the folate-dependent strain, pABA and folate for creating standard assay curves, and the complete *L.lactis* genome.

## WORKS CITED

Asrar, Farhan M., and Deborah L. O'Connor. "Bacterially synthesized folate and supplemental folic acid are absorbed across the large intestine of piglets." *Journal of Nutritional Biochemistry* , 2005: 16: 587–593.

Bermingham, Alun, and Jeremy P. Derrick. "The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery." *BioEssays* , 2002: 24:637–648.

Bolotin, Alexander, et al. "The Complete Genome Sequence of the Lactic Acid Bacterium Lactococcus lactis ssp. lactis IL1403." *Genome Res.*, 2001: 11: 731-753.

CAMILO, ERMALINDA, JOSEPH ZIMMERMAN, JOEL B. MASONl, and BARBARA GOLNER. "Folate Synthesized by Bacteria in the Human Upper Small Intestine Is Assimilated by the Host." *GASTROENTEROLOGY* , 1996: 110:991–998.

Crécy-Lagard, Valérie de, Basma El Yacoubi, and Rocío Díaz de la Garza. "Comparative genomics of bacterial and plant folate synthesis and salvage: predictions and validations." *BMC Genomics* , 2007: 8:245.

Endy, Drew. "Foundations for engineering biology." *Nature*, 2005: 449-453.

Horne, Donald W., and Dewaln Patterson. "Lactobacillus casei Microbiological Assay of Folic Acid Derivatives in 96-Well Microtiter Plates." *CLIN. CHEM.* , 1988: 34/11, 2357-2359 .

Sybesma, WFH. "Metabolic Engineering of Folate Production in Lactic Acid Bacteria." (Thesis Wageningen University) 2003.

Sybesma, Wilbert, Marjo Starrenburg, Michiel Kleerebezem, and Igor Mierau. "Increased Production of Folate by Metabolic Engineering of Lactococcus lactis." *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, June 2003: 3069–3076.

Wegkamp, Arno, Marjo Starrenburgl, Willem M. de Vos, Jeroen Hugenholtz, and Wilbert Sybesma. "Transformation of Folate-Consuming Lactobacillus gasseri into a." *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, May 2004: 3146–3148.

Wegkamp, Arno, Wietske van Oorschot, Willem M. de Vos, and Eddy J. Smid. "Characterization of the Role of para-Aminobenzoic Acid Biosynthesis." *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, Apr. 2007: 2673–2681.

Wikipedia. *Gut Flora.* http://en.wikipedia.org/wiki/Gut_flora (accessed June 2008).

Zhu, T., Z. Pan, N. Domagalski, R. Koepsel, M. M. Ataai, and M. M. Domach. "Engineering of Bacillus subtilis for Enhanced Total Synthesis." *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, Nov. 2005: 7122–7129.